|  |
| --- |
|  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Chi-Square (X2) Basics**  Ms. Ottolini, AP Biology, 2013-2014  **When do I use a Chi square (X2) test?**   * The X2 test is a statistical test to compare observed results with theoretical expected results. * The calculation generates a X2 value; the higher the value of X2, the greater the difference between the observed and the expected results.   **How do I perform a Chi square test?**   1. **State the null hypothesis**  * This is a negative statement, basically saying that there is no statistically significant difference between the observed and the expected results or between two sets of data.  1. **Determine your expected values**  * The way you calculate your expected values will be different for each situation * For example, if your null hypothesis states that there are no differences between the observed genotype frequencies in a population and the genotype frequencies predicted by Hardy-Weinberg equilibrium, you will need to find p2, q2, and 2pq to determine your expected values. * Expected and observed values are always whole numbers. Therefore, if you want to find the number of individuals expected for the homozygous dominant genotype, you will multiply p2 by the total population size to determine the number of individuals expected for that genotype.   **3. Calculate chi2**   * The formula is: * Where o = observed value, e = expected value, and ∑ = the sum of * So you would need to calculate separately for each value (ex: each genotype) and then add the results together   **4. You will also need to know the degrees of freedom.**   * This is calculated using the formula (n-1) * where n = the number of sets of results.(ex: the number of possible genotypes) * For the Hardy-Weinberg situation above… degrees of freedom = n-1 = 3-1 = 2   **5. Compare the X2 value against a table of critical values.**   * On the table below, refer to the row that corresponds to the correct number of degrees of freedom for your data set * Look up the critical number at the p = 0.05 level. “p” stands for probability     **6. Make a conclusion**   * If the X2 value that you calculated in Step 3 is higher than the critical number at the p = 0.05 level then you can reject the null hypothesis. In other words, there is a statistically significant difference between the observed and expected results. (i.e. the observed results do not match the expected results)   *Note: A high X2 value corresponds with a low p value (below 0.05)*   * If the X2 value is less than the critical number then you can accept the null hypothesis. In other words, there is no statistically significant difference between the observed and expected results. (i.e. the observed results match the expected results)   *Note: A low X2 value corresponds with a high p value (above 0.05)*  **Chi Square Test: Example #1**  Naked mole rats are a burrowing rodent native to parts of East Africa[.](http://en.wikipedia.org/wiki/East_Africa) They have a complex social structure in which only one female (the queen) and one to three males reproduce, while the rest of the members of the colony function as workers. Mammal ecologists suspected that they had an unusual male to female ratio. They counted the numbers of each sex in one colony.   |  |  | | --- | --- | | Sex | Number of animals | | Female | 52 | | Male | 34 |   **State the Null hypothesis**  **Calculate the expected results**  **Calculate the chi-squared value**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Sex | Observed | Expected | O - E | (O – E)2 | (O – E)2/E | | Female | 52 |  |  |  |  | | Male | 34 |  |  |  |  | | TOTAL |  |  |  |  |  |   X2 =  **What are the degrees of freedom?**  DF =  **Compare the calculated value with the critical value**    **Make a conclusion** |
|  |

**Chi Square Test: Example #2**

You have been wandering about on a seashore and you have noticed that a small snail (the flat periwinkle) seems to live only on seaweeds of various kinds. You decide to investigate whether the animals prefer certain kinds of seaweed by counting numbers of animals on different species. You end up with the following data

|  |  |
| --- | --- |
| TYPE OF SEAWEED | Number of animals on each kind of seaweed |
| serrated wrack | 45 |
| bladder wrack | 38 |
| egg wrack | 10 |
| spiral wrack | 5 |
| other algae | 2 |
| TOTAL | 100 |

**State the Null hypothesis**

**Calculate the expected results**

**Calculate the chi-squared value**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Seaweed | Observed | Expected | O - E | (O – E)2 | (O – E)2/E |
| serrated wrack | 45 |  |  |  |  |
| bladder wrack | 38 |  |  |  |  |
| egg wrack | 10 |  |  |  |  |
| spiral wrack | 5 |  |  |  |  |
| other algae | 2 |  |  |  |  |
| TOTAL |  |  |  |  |  |

X2 =

**What are the degrees of freedom?**

DF =

**Compare the calculated value with the critical value**

****

**Make a conclusion**

**Chi Square Test: Example #3 (The Hardy Weinberg Equilibrium Lab)**

For **Case #1** of your Hardy-Weinberg Equilibrium Lab, determine if the class observed values for the genotype frequencies in **generation F5** match with the genotype frequencies you would expect for a population in Hardy Weinberg equilibrium. Remember, the initial p and q values are both 0.5

**State the Null hypothesis**

**Calculate the expected results**

Remember, your expected results must be whole numbers (not decimals). Therefore to find the number of individuals expected for each genotype in THIS POPULATION, multiply the frequency for a particular genotype (ex: p2 for the homozygous genotype) by the total number of individuals in the population.

|  |  |
| --- | --- |
| # of individuals expected to have the homozygous dominant genotype (AA) | p2 x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |
| # of individuals expected to have the heterozygous genotype (Aa) | 2pq x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |
| # of individuals expected to have the homozygous recessive genotype (aa) | q2 x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |

**Calculate the chi-squared value**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype** | **Observed** | **Expected** | **O - E** | **(O – E)2** | **(O – E)2/E** |
| AA |  |  |  |  |  |
| Aa |  |  |  |  |  |
| aa |  |  |  |  |  |
| Total |  |  |  |  |  |

X2 =

**What are the degrees of freedom?**

DF =

**Compare the calculated value with the critical value**



**Make a conclusion**

**Chi Square Test: Example #4 (The Hardy Weinberg Equilibrium Lab)**

For **Case #2** of your Hardy-Weinberg Equilibrium Lab, determine if the class observed values for the genotype frequencies in **generation F5** match with the genotype frequencies you would expect for a population in Hardy Weinberg equilibrium. Remember, the initial p and q values are both 0.5

**State the Null hypothesis**

**Calculate the expected results**

Remember, your expected results must be whole numbers (not decimals). Therefore to find the number of individuals expected for each genotype in THIS POPULATION, multiply the frequency for a particular genotype (ex: p2 for the homozygous genotype) by the total number of individuals in the population.

|  |  |
| --- | --- |
| # of individuals expected to have the homozygous dominant genotype (AA) | p2 x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |
| # of individuals expected to have the heterozygous genotype (Aa) | 2pq x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |
| # of individuals expected to have the homozygous recessive genotype (aa) | q2 x (total # of individuals in the population) = \_\_\_\_\_\_ x \_\_\_\_\_\_ = \_\_\_\_\_\_ |

**Calculate the chi-squared value**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype** | **Observed** | **Expected** | **O - E** | **(O – E)2** | **(O – E)2/E** |
| AA |  |  |  |  |  |
| Aa |  |  |  |  |  |
| aa |  |  |  |  |  |
| Total |  |  |  |  |  |

X2 =

**What are the degrees of freedom?**

DF =

**Compare the calculated value with the critical value**



**Make a conclusion**